Statistics 581, Problem Set 10

Wellner; 11/28/2018

Reading: Course Notes, Chapter 4, Sections 1-4;
Ferguson, ACLST, Chapters 20, Chapter 22, and Chapter 16
vdV, Asymp. Statist., sections 5.6 and 5.7, pages 67 - 75.

Due: Wednesday, December 5, 2018.
Reminder: Final Exam; Monday, December 10, 2018: 8:30-10:20, DEN 112

1. Ferguson, ACLST, page 150, problem 3. Does the theory in our Chapter 4 (or Ferguson's Chapter 22) apply directly? Does the local asymptotic power of your test depend on the common value of θ_j in the null hypothesis?

2. Ferguson, ACLST, page 149, problem 2 modified as follows:
(a) Find the LR test statistic of the null hypothesis $H_0 : \mu = c\theta$ for any fixed number $c > 0$, and find the asymptotic distribution of the LR statistic under H_0.
(b) Does the theory of our chapter 4 (or Ferguson's chapter 22) apply directly?
(c) Does the local asymptotic power of your test depend on c?

3. Ferguson, ACLST, page 118, problem 3. (See also Example 4.3.7, page 21, Chapter 4 notes.) [Neyman and Scott (1948)] Suppose we have a sample of size d from each of n normal populations with common unknown variance but possibly different unknown means $X_{i,j} \sim N(\mu_i, \sigma^2)$, $i = 1, \ldots, n$, $j = 1, \ldots, d$ where all the $X_{i,j}$ are independent.
(a) Find the maximum-likelihood estimate of σ^2.
(b) Show that for d fixed the MLE of σ^2 is not consistent as $n \to \infty$. Why don't either of Theorem 17 (Ferguson) or our Theorem 4.1.2 apply?
(c) Find a consistent estimate of σ^2.

4. Consider the Weibull family of example 3.2.5 and problem set #6, problem 1: $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ with $\Theta \subset R^{+2}$ given by the (Lebesgue) densities

$$p_\theta(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) 1_{[0,\infty)}(x)$$

where $\theta \equiv (\alpha, \beta) \in (0, \infty) \times (0, \infty) \subset R^2$. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with density function p_θ.
(a) If $X \sim P_\theta \in \mathcal{P}$, show that the distributions of $\log X$ form a location and scale family from a Gumbel (extreme value) density on R. (This amounts to a rephrasing of the statement of a problem in an earlier problem set.)
(b) Use the result of (a) to construct method of moments estimators or quantile based estimators $\overline{\theta}_n$ of $\theta = (\alpha, \beta)$.

(c) Show that the method of moments or quantile estimators $\overline{\theta}_n$ of θ are asymptotically normal, and find the asymptotic distribution; i.e. show that

$$\sqrt{n}(\overline{\theta}_n - \theta) \to_d N_2(0, \Sigma) \qquad \text{for some} \qquad \Sigma.$$

[We will use these estimators as "starting points" approximate (or one-step) maximum likelihood estimators in the next problem .]

5. (Problem 4, continued).
(a) Does a maximum likelihood estimate of $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ exist? Is it unique? (See Lehmann and Casella, Example 6.1, page 468.)
(b) Compute an approximate (one - step) maximum likelihood estimate $\check{\theta}$ of θ using the method of moment (or quantile) estimators $\overline{\theta}_n$ as the preliminary estimators based on the following data (with $n = 12$):

$$1,\ 1,\ 2,\ 3,\ 14,\ 27,\ 41,\ 55,\ 66,\ 113,\ 320,\ 413.$$

[These are failure times in seconds for "breakdown" of an insulating fluid between two electrodes subject to a voltage of 40 kV. – from Nelson, *Applied Life Data Analysis*, page 252, modified slightly.]
(c) Compute the maximum likelihood estimator $\hat{\theta}_n$, and compare it with the one step estimator computed in (b).

6. **Optional bonus problem 1:**
(a) Ferguson, ACLST, page 139, problem 3.
(b) What if Ferguson's density $f(x|\theta)$ with $\theta \in (0,1)$ is replaced by $\theta = (\gamma, \eta) \in (0,1) \times (0, \infty)$ and

$$f(x|\theta) \equiv f(x|\gamma, \eta) = \{(1-\gamma)e^{-x} + \gamma\eta^2 x \exp(-\eta x)\} 1_{[0,\infty)}(x)?$$

Can you estimate γ and η by the method of moments? Can you improve method of moment estimators via one-step estimators?

7. **Optional bonus problem 2:** Suppose that (as in Lemma 5.2, page 38, Chapter 3 Notes) P and Q are two probability measures on a measurable space $(\mathcal{X}, \mathcal{A})$ with densities p and q with respect to a $\sigma-$finite dominating measure μ, and P^n and Q^n denote the corresponding product measures on $(\mathcal{X}^n, \mathcal{A}_n)$ (of $X_1, \ldots, X_n$ i.i.d. as P or Q respectively).
(a) What is the relationship between $K(P^n, Q^n)$ and $K(P, Q)$, if any?
(b) If P is the Normal$(0, \sigma^2)$ distribution and Q is the Normal(μ, σ^2) distribution, compute $K(P, Q)$, $\rho(P, Q) = \int \sqrt{pq}\, d\mu$, and $H^2(P, Q)$.
(c) Use the results of (a) and (b) together with Lemma 5.2 to calculate $K(P^n, Q^n)$, $\rho(P^n, Q^n)$, and $H^2(P^n, Q^n)$ when P and Q are as in (b).
(d) Find a sequence μ_n so that, with Q_n being the Normal distribution with mean μ_n, the quantities $K(P^n, Q_n^n)$, $\rho(P^n, Q_n^n)$, and $H^2(P^n, Q_n^n)$ converge to finite limits as $n \to \infty$.

8. **Optional bonus problem 3:**
(a) Prove the following inequality relating the Hellinger distance $H^2(P,Q)$ to $K(P,Q)$: $2H^2(P,Q) \le K(P,Q)$. Since $d_{TV}(P,Q) \le \sqrt{2}H(P,Q)$ this implies that $d_{TV}(P,Q) \le \sqrt{K(P,Q)}$. **Hint:** Start with $K(P,Q)$ and use the inequality $-\log(1+x) \ge -x$ for all $x > -1$.
(b) Suppose that P and Q have densities p and q with respect to a common dominating measure μ. Show that $K(P,Q) = \int_{pq>0} p \log(p/q) d\mu$. (c) Let $h(x) \equiv x \log x - x + 1$ for $x \ge 0$ with $h(0) = 1$. show that $h(1) = 0$, $h'(1) = 0$, $h''(x) = 1/x \ge 0$, and $h(x) \ge 0$ for all $x \ge 0$. Moreover, show that

$$\frac{2}{3}(2+x)h(x) \ge (x-1)^2 \quad \text{for all } x \ge 1.$$

(d) Use the inequality in the display above to prove Pinsker's inequality: $d_{TV}(P,Q) \le \sqrt{K(P,Q)/2}$. **Hint:** Note that if $P \prec Q$ (P is absolutely continuous with respect to Q), then $d_{TV}(P,Q) = \int_{q>0} |(p/q) - 1| q d\mu$; then use the inequality in (a) together with the Cauchy-Schwarz inequality. The inequality in (a) can be rewritten as $\psi(x) \ge 1/(1 + x/3)$ where $\psi(x) \equiv (2/x^2)h(1+x)$ which arises in exponential bounds for the Binomial distribution; see e.g. Shorack & W (1986, 2009) Proposition 11.1.1, page 441.